2012 27th Annual ACM/IEEE Symposium on Logic in Computer Science

Strong Complementarity and Non-locality in Categorical Quantum Mechanics

Bob Coecke[1] Ross Duncan[2] Aleks Kissinger[1] Quanlong Wang[3]

[1]University of Oxford, Department of Computer Science,
Wolfson Building, Parks Road, Oxford OX1 3QD, UK.
coecke/alek@cs.ox.ac.uk
[2]Université Libre de Bruxelles, Laboratoire d'Information Quantique
Campus Plaine, Boulevard du Triomphe, 1050 Brussels, Belgium
rduncan@ulb.ac.be
[3]Beihang University, School of Mathematics and System Sciences
XueYuan Road No.37, HaiDian District, Beijing, China
qlwang@buaa.edu.cn

***Abstract*—Categorical quantum mechanics studies quantum theory in the framework of dagger-compact closed categories.**

Using this framework, we establish a tight relationship between two key quantum theoretical notions: non-locality and complementarity. In particular, we establish a direct connection between Mermin-type non-locality scenarios, which we generalise to an arbitrary number of parties, using systems of arbitrary dimension, and performing arbitrary measurements, and, a new stronger notion of complementarity which we introduce here.

Our derivation of the fact that strong complementarity is a necessary condition for a Mermin scenario provides a crisp operational interpretation for strong complementarity. We also provide a complete classification of strongly complementary observables for quantum theory, something which has not yet been achieved for ordinary complementarity.

Since our main results are expressed in the (diagrammatic) language of dagger-compact categories, they can be applied outside of quantum theory, in any setting which supports the purely algebraic notion of strongly complementary observables. We have therefore introduced a method for discussing non-locality in a wide variety of models in addition to quantum theory.

The diagrammatic calculus substantially simplifies (and sometimes even trivialises) many of the derivations, and provides new insights. In particular, the diagrammatic computation of correlations clearly shows how local measurements interact to yield a global overall effect. In other words, we *depict non-locality*.

***Index Terms*—quantum computing, quantum entanglement, abstract algebra**

I. Introduction

This paper is concerned with two central notions in quantum foundations and quantum computation, *non-locality* and *complementarity*, and the study of their relationship.

Non-locality is what Einstein notoriously referred to as 'spooky action at a distance'. It was formally substantiated for the first time by Bell's theorem, and experimentally verified by testing Bell-inequalities. It states that the correlations observed when measuring spatially quantum separated systems cannot be explained by means of classical probabilities, i.e. that there is no underlying *hidden variable theory*. Complementarity, informally put, states that if one knows the value of one observable sharply (e.g. position), then there is complete uncertainty about the other observable (e.g. momentum).

These two concepts underpin what is arguably the most successful endeavour towards quantum information technologies: quantum cryptography. Indeed, in a quantum key distribution protocol, encoding data in either of two complementary observables will enable the parties to detect interception by an adversary [5], while non-locality allows one to verify the authenticity of the shared entangled resource by means of which key sharing is established [20].

This research applies methods of computer science and logic to investigations in quantum foundations. It is moreover strongly aligned with the current trend in the broader quantum information community: understanding quantum information processing within a larger space of hypothetical information processing theories in order to understand what is particular about quantum theory.

Until now, this area has been characterised by the study of *generalised probabilistic theories* [4], a space of theories which includes quantum probability theory, classical probability theory, as well as theories which are even more non-local than quantum theory. A topic of particular focus has been the search for peculiarities of quantum non-locality, within the larger space of non-local theories, e.g. [32], [3].

Of course, a space of more general theories can be conceived as abstracting away certain concrete features of quantum theory, hence encompassing a broader class of theories, and the words 'generalised' and 'abstract' can be treated as synonymous. While the generalised probabilistic theories discussed above abstract away all but convex probabilistic structure, our focus is on the *compositional structure on processes*, say *generalised process theories*.

While having composition play a leading role evidently draws form computer science practice, it also appeals to Schrödinger's conviction that what mostly characterises quantum theory is the manner in which systems compose [33].

1043-6871/12 $26.00 © 2012 IEEE
DOI 10.1109/LICS.2012.35

This compositional paradigm was the main motivation for *categorical quantum mechanics* (CQM), initiated by Abramsky and BC in [2]. Meanwhile, CQM has helped to solve open problems in quantum information and computation e.g. [12], [17], [26], and has meanwhile been adopted by leading researchers in the area of quantum foundations e.g. [7], [22].

In this paper we push CQM far beyond its previously established horizons, both on the level of comprehensiveness and in terms of its application domain. The vehicle to do so is Mermin's ingenious but elaborate non-locality proof [6], [23], [21], [29]. This argument, usually stated mostly in natural language, establishes non-locality as a *contradiction of parities* for quantum theory and local theories, rather than as the violation of a Bell-inequality, and involves a sophisticated interplay of several incompatible measurement scenarios, each involving measurements against varying angles, as well as hypothetical hidden variable theories, and the manipulations of the resulting probabilistic measurement data.

Our abstraction as well as our generalisations of this scenario provides important new physical insights in the nature of non-locality, and its relationship to complementarity, as we discuss in Section VIII. Hence this paper is both one in computer science as well as one in quantum foundations.

Outline. Sec. II provides an overview of CQM and presents some relevant non-standard models. We briefly recall the diagrammatic calculus for symmetric monoidal categories.

In Sec. III we define strong complementarity, establish its relationship with ordinary complementarity, and state the first major result of this paper: the classification of strongly complementary observables in quantum theory.

In Sec. IV we diagrammatically compute the correlations of measurements against arbitrary angles on an n-party GHZ-state for systems of arbitrary type.

In Sec. V we cast the usual Mermin argument within CQM, as a stepping stone to its generalisation in Sec. VI. In doing so we rely crucially on strong complementarity.

Finally, in Sec. VII we establish the necessity of strong complementarity for Mermin arguments. This also provides an operational interpretation for strong complementarity.

Due to space restrictions, many proofs have been omitted. These, together with additional background material, may be found in the preprint version of this paper available from http://arxiv.org/abs/1203.4988.

Earlier Work. In [8], [9] BC and RD introduced the graphical ZX-*calculus* for the specific purpose of reasoning graphically about qubits. This calculus included the equations that we identify here as strong complementarity, but these were not identified as such. In [11] BC, Edwards and Spekkens relied on the CQM concept of a phase groups to identify the differences of the categories $\mathbf{Stab}$ and $\mathbf{Spek}$ with respect to non-locality, but the Mermin scenario was not formulated within CQM. In [18], [19] Edwards expanded this line work and derived some higher-dimensional generalisations of the Mermin argument.

II. Background I

A. Models of Physical Theories

Symmetric monoidal categories (SMCs) provide a very general language for physical theories: a morphism $f : A \to B$ is interpreted as a process from a physical system of type A to one of type B. Spatial and temporal extension are represented by the composition and tensor product; that is, by the sequential and parallel combination of processes. A *state* of system A is simply a morphism $\psi : I \to A$, while an *effect* has type $\pi : A \to I$. Every SMC has a commutative monoid of *scalars*, the morphims $s : I \to I$.

Each concrete SMC is a model of this primitive theory, with different physical characteristics. For example, $(\mathbf{FHilb}, \otimes)$, the category of of finite dimensional Hilbert spaces and linear maps, comprises quantum systems and pure post-selected quantum processes[1]. We write $\mathbf{FHilb}_D$ to denote the category of Hilbert spaces of dimension D^n, for some fixed D. In particular, $\mathbf{Qubit} := \mathbf{FHilb}_2$. In fact, $\mathbf{FHilb}$ has redundancies: two linear maps that are equal up to a non-zero scalar multiple represent the same physical process.

This redundancy can be eliminated by using a different model. The category $CP(\mathbf{FHilb})$ models open quantum systems and completely positive maps. It is built from $\mathbf{FHilb}$ using Selinger's CP-construction [34], described Sec. II-E, which constructs a category of *mixed* processes from any given category of pure processes.

Despite its seemingly different nature, $(\mathbf{FRel}, \times)$, the category of finite sets and relations with the Cartesian product as the tensor is a key example. Restricting to powers of a fixed set of size D we write $\mathbf{FRel}_D$.

In fact, $\mathbf{FRel}$ shares many features of $\mathbf{FHilb}$ and $CP(\mathbf{FHilb})$. All three are *compact* categories [27]: each object A has a dual object A^* and morphisms $\eta_A : I \to A^* \otimes A$ and $\varepsilon_A : A \otimes A^* \to I$ such that $(\varepsilon_A \otimes 1_A) \circ (1_A \otimes \eta_A) = 1_A$ and $(1_{A^*} \otimes \varepsilon_A) \circ (\eta_A \otimes 1_{A^*}) = 1_{A^*}$. We refer to these elements as the *compact structure* on A. Note that they need not be unique: an object may support several compact structures.

Further, our three examples are all *dagger compact* [2]: there exists an identity-on-objects involutive contravariant strict monoidal endofunctor $\dagger : \mathbf{C} \to \mathbf{C}$ with $\varepsilon_A = \eta_A^\dagger \circ \sigma_{A^*,A}$ for all objects A, to which we refer as the *dagger functor*.

Dagger compactness provides the language for many quantum concepts such as Bell-states/effects, unitarity, the Born rule [2], and complete positivity [34]. E.g. *unitarity* means $U^\dagger \circ U = 1_A$ and $U \circ U^\dagger = 1_B$. An SMC with a dagger functor is referred to as $\dagger$-SMC. In any $\dagger$-SMC we can define an inner product for states ψ and ϕ as $\psi^\dagger \circ \phi : I \to I$. This provides the usual inner-product (in $\mathbf{FHilb}$) as well as amplitudes (in $CP(\mathbf{FHilb})$).

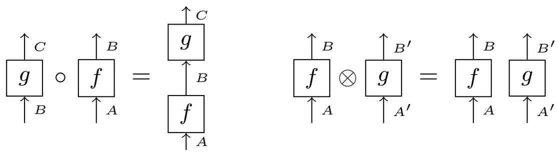
B. $\dagger$-Compact Categories and Diagrammatic Calculus

Monoidal categories admit a diagrammatic notation which greatly simplifies the task of reading, analysing, and comput-

[1] 'Pure' as in 'not mixed'; 'post-selected' means that processes which are conditional on obtaining a given measurement outcome are permitted.

ing in this framework. For the †-compact categories of interest here this language takes a particularly simple form.

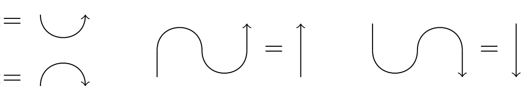
Systems (objects) are depicted by labelled wires, and processes (morphisms) are represented as boxes with wires in and out, indicating the type of the process. Composition is expressed by plugging the outputs of one box into the inputs of another, and the monoidal product is given by juxtaposition. The monoidal unit I is represented as the empty diagram.



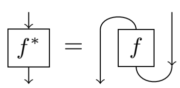
We shall draw composition from bottom-to-top, and omit wire labels where there is no ambiguity. In a symmetric monoidal category, we indicate the swap map as a wire crossing. For compact closed categories, we indicate a dual object A^* by a wire of the opposite direction. The cap and cup maps are then depicted as half-turns.

$e_A :=$

$d_A :=$



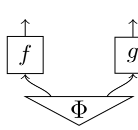
Using caps and cups, we can turn any morphism $f : A \to B$ into a morphism on the dual objects going in the opposite direction: $f^* : B^* \to A^*$.



This is sometimes called the transpose of f, but this terminology can be misleading. In **FHilb**, L^* is the map that takes a linear form $\langle\xi| \in B^*$ to $\langle\xi|L \in A^*$. We refer to this map simply as the *upper-star* of f. In a †-category, the † functor sends $f : A \to B$ to $f^\dagger : B \to A$. In a †-compact category, we define the *lower-star* of f as $f_* := (f^\dagger)^* = (f^*)^\dagger$.

Since our category is symmetric, wires are allowed to cross, and boxes may slide up along wires freely without changing the denotation of the diagram. More generally, if one diagram may be deformed continuously to another, then these diagrams denote the same arrow in the category. For a more comprehensive description of graphical languages, see Selinger [35].

In the diagrams to come, we will often use horizontal separation to indicate separation in space and vertical separation to indicate separation in time. For example,



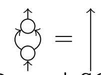
depicts the creation of two systems by the process Φ, which then become spatially separated over some time and are acted upon by processes f and g respectively.

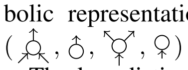
C. Generalised Observables

An *observable* yields classical data from a physical system [14], [13]. In quantum mechanics, an observable is represented by a self-adjoint operator. The important information encoded by a (non-degenerate) observable is its orthonormal basis of eigenstates. In **FHilb**, orthonormal bases (ONBs) are in 1-to-1 correspondence with †-special commutative Frobenius algebras [15].

Definition 2.1: In a †-SMC, a †-special Frobenius algebra (†-SCFA) is a commutative Frobenius algebra

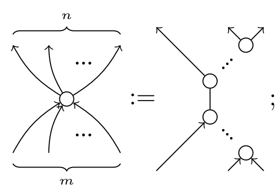
$$\mathcal{O}_\circ = (\mu_\circ : X \otimes X \to X, \quad \eta_\circ : I \to X, \\ \delta_\circ : X \to X \otimes X, \quad \epsilon_\circ : X \to I)$$

such that $\delta_\circ = (\mu_\circ)^\dagger$, $\epsilon_\circ = (\eta_\circ)^\dagger$ and .

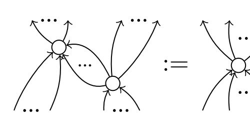
Because the correspondence to ONBs, a †-SCFA is also called an *observable structure*. We will use the symbolic representation $(\mu_\circ, \eta_\circ, \delta_\circ, \epsilon_\circ)$ and the pictorial one () interchangeably.

The key distinction between classical and quantum data is that classical data may be freely copied and deleted while this is impossible for quantum data, due to the no-cloning [16], [38] and no-deleting [30] theorems. This is a general fact: any compact closed category with natural diagonal maps $\Delta : X \to X \otimes X$ collapses to a preorder [1].

Proposition 2.2: Given an observable structure $\mathcal{O}_\circ$ on X, let $(\bigcirc)_n^m$ denote the '(n, m)-legged spider':



then any morphism $X^{\otimes n} \to X^{\otimes m}$ built from $\mu_\circ, \eta_\circ, \delta_\circ$ and $\epsilon_\circ$ via †-SMC structure which has a connected graph is equal to the $(\bigcirc)_n^m$. Hence, spiders compose as follows:

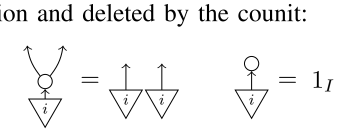


(1)

Concretely, given an ONB $\{|i\rangle\}_i$ then $\delta_\circ :: |i\rangle \mapsto |ii\rangle$ defines an observable, and all observables are of this form for some ONB. The resulting intuition is that $\delta_\circ$ is an operation that 'copies' basis vectors, and that $\epsilon_\circ$ 'erases' them [13].

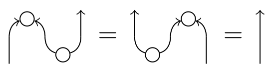
Perhaps surprisingly, **FRel** also has many distinct observables, which have been classified by Pavlovic [31]. Even on the two element set there are two [10], namely $\delta_\circ : \{0,1\} \to \{0,1\} \times \{0,1\} :: i \mapsto (i,i)$ and $\delta_\circledcirc :: 0 \mapsto \{(0,0),(1,1)\}; 1 \mapsto \{(0,1),(1,0)\}$. In fact, this pair is strongly complementary in the sense of Sec. III.

Each observable structure comes with a set of *classical points*, the abstract analogues to eigenvectors of an observables. A classical point is a state that is copied by its comultiplication and deleted by the counit:



(2)

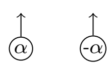
We will depict classical points as triangles of the same colour as their observable structure. Each observable structure furthermore defines a *self-dual* †-compact structure on its object. That is, it defines a †-compact structure where the dual object of X is X itself. By (1) we have:



The upper-star with respect to this compact structure corresponds in **FHilb** to transposition in the given basis. For that reason, we call this the $\bigcirc$-transpose $f^{\circledast}$. The lower star corresponds to complex conjugation in the basis of $\mathcal{O}_\circ$, so we call it the $\bigcirc$-conjugate $f_{\circledast} := (f^{\circledast})^\dagger$.

D. Phase Group for an Observable Structure

Given an observable structure $\mathcal{O}_\circ$ on X, the multiplication $\mu_\circ$ puts a monoid structure on the points of X. If we restrict to those points $\psi_\alpha : I \to A$ where $\mu_\circ(\psi_\alpha \otimes (\psi_\alpha)_{\circledast}) = \eta_\circ$, we obtain an Abelian group $\Phi_\circ$ called the *phase group* of $\mathcal{O}_\circ$ [8]. We let $\psi_{-\alpha} := (\psi_\alpha)_{\circledast}$ and represent these points as circles with one output, labelled by a phase.

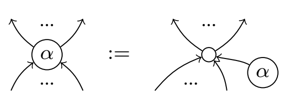


For an observable on the qubit, the phase group consists of the points that are unbiased to the eigenstates of the observable, and their multiplication is their convolution. Explicitly, given $\mu_\circ = |0\rangle\langle 00| + |1\rangle\langle 11|$, we have:

$$\mu_\circ \left(\begin{pmatrix} 1 \\ e^{i\alpha} \end{pmatrix} \otimes \begin{pmatrix} 1 \\ e^{i\beta} \end{pmatrix} \right) = \begin{pmatrix} 1 \\ e^{i(\alpha+\beta)} \end{pmatrix},$$

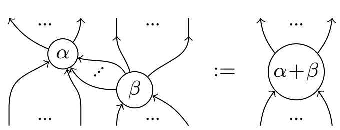
so we obtain the circle group.

We can now introduce 'spiders decorated with phases':



(3)

which compose as follows:



(4)

The name 'phase group' comes from fact that phased spiders with one input and one output are the abstract analogue to phase gates, familiar from quantum computing. This is supported by the following fact, proven in [9].

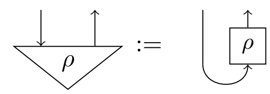
Proposition 2.3: If $\psi_\alpha \in \Phi_\circ$ then $\mu_\circ \circ (1_X \otimes \psi_\alpha) : X \to X$ (i.e. a spider with one input and one output, labelled by α) is unitary.

E. Generalised Classical-Quantum Interaction

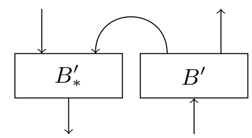
In ordinary quantum theory, quantum states are represented as positive operators and operations as completely positive maps, or CPMs. These are maps that take positive operators to positive operators. A general CPM can be written in terms of a set of linear maps $\{B_i : \mathcal{H}_1 \to \mathcal{H}_2\}$ called its *Kraus maps*.

$$\Theta(\rho) = \sum_i B_i \rho B_i^\dagger$$

Since **FHilb** is compact, we can regard $\rho : \mathcal{H}_1 \to \mathcal{H}_1$ as an element of $\mathcal{H}_1^* \otimes \mathcal{H}_1$.

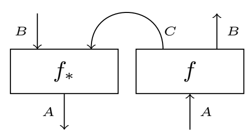


Then, we can encode the Kraus vectors of Θ in a map $B' = \sum |i\rangle \otimes B_i$ and represent Θ as:



(5)

When we take the elements in Eq. (6) to be morphisms in an arbitrary †-compact category, this gives us an abstract definition of a completely positive map. This is Selinger's representation of CPMs [34].



(6)

Important special cases are *states* where $A \cong I$, *effects* where $B \cong I$, and 'pure' maps, where $C \cong I$.

Returning to quantum mechanics, we can see how a quantum measurement would look in this language. A (projective) quantum measurement $M_\circ$ is a CPM that sends trace 1 positive operators (in this case quantum states) to trace 1 positive operators that are diagonal in some ONB (encoding a probability distribution of outcomes). Suppose we wish to measure with respect to $\mathcal{O}_\circ$, whose classical points form an ONB $\{|x_i\rangle\}$. The probability of getting the i-th measurement outcome is computed using the Born rule.

$$\text{Prob}(i, \rho) = \text{Tr}(|x_i\rangle\langle x_i| \rho) = \langle x_i| \rho |x_i\rangle$$

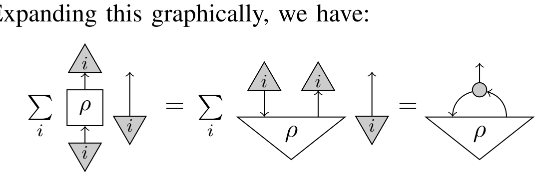
We can encode this map from states to distributions as:

$$M_\circ(\rho) = \sum_i (\langle x_i| \rho |x_i\rangle) |x_i\rangle\langle x_i|$$

Or, rather than encoding the distribution in a diagonal matrix, we could simply use a vector:

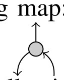
$$m_\circ(\rho) = \sum_i (\langle x_i| \rho |x_i\rangle) |x_i\rangle$$

Expanding this graphically, we have:

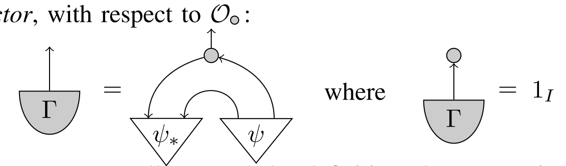


We are now ready to make definitions for abstract measurements and abstract probability distributions, which we shall call Born vector.

Definition 2.4: For an observable structure $\mathcal{O}_{\bullet}$, a measurement is defined as the following map:

$m_{\bullet} :=$ 

Any point $|\Gamma) : I \to X$ of the following form is called a *Born vector*, with respect to $\mathcal{O}_{\bullet}$:

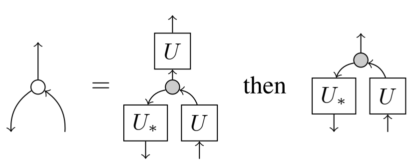


We can naturally extend the definition above to points of the form $|\Gamma) : I \to X \otimes \ldots \otimes X$ by requiring that they be Born vectors with respect to the product Frobenius algebra $\mathcal{O}_{\bullet} \otimes \ldots \otimes \mathcal{O}_{\bullet}$.

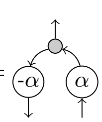
The adjoint of the measurement map $m_{\bullet}^{\dagger}$ is a preparation operation. In **FHilb**, it takes a Born vector $|\Gamma)$ with respect to $\mathcal{O}_{\bullet}$ and produces a probabilistic mixture of the (pure) outcome states of $\mathcal{O}_{\bullet}$ with probabilities given by $|\Gamma)$.

This leads to a simple classical vs. quantum diagrammatic paradigm that applies to arbitrary observables in any †-SMC [13]: *classical systems are encoded as a single wire and quantum systems as a double wire*. The same applies to operations, and $m_{\bullet}$ and $m_{\bullet}^{\dagger}$ allow passage between these types.

Note that the classical data will 'remember' to which observable it relates, cf. the encoding $\sum_i p_i |x_i\rangle$. This is physically meaningful since, for example, when one measures position the resulting value will carry specification of the length unit in which it is expressed. If one wishes to avoid interconversion of this 'classical data with memory', one could fix one observable, and unitarily transform the quantum data before measuring. Indeed, if



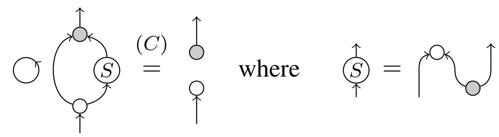
measures the $\mathcal{O}_{\circ}$-observable but produces $\mathcal{O}_{\bullet}$-data. In **FHilb**, all observable structures are unitarily isomorphic, so any projective measurement can be obtained in this way. A particularly relevant example is when these unitaries are phases with respect the another observable structure $\mathcal{O}_{\circ}$.

$$m_{\bullet}^{\alpha} :=$$  (7)

When $\mathcal{O}_{\circ}$ is induced by the Pauli spin-Z observable and $\mathcal{O}_{\bullet}$ by the Pauli spin-X observable, then $m_{\bullet} = m_{\bullet}^{0}$ is an X measurement and $m_{\bullet}^{\pi/2}$ is a Y measurement. Note however, that both produce Born vectors of outcome probabilities with respect to the ○ basis. This will be useful in the sections to come.

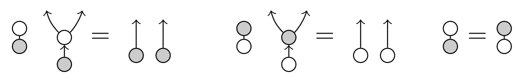
III. Strong Complementarity

Definition 3.1: A pair $(\mathcal{O}_{\circ}, \mathcal{O}_{\bullet})$ of observables on the same object X is *complementary* iff:

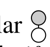


If at least one of the two observables has 'enough classical points', this equation holds if and only if the classical points of one observable are 'unbiased' (in sense of [8]) for the other observable. Every observable in **FHilb** has enough classical points, hence we reclaim the usual notion of quantum complementarity, and extend it to a more general setting.

Definition 3.2: A pair $(\mathcal{O}_{\circ}, \mathcal{O}_{\bullet})$ of observables on the same object X is *coherent* iff:



In other words, $\epsilon_{\circ}$ is proportional to a classical point for $\mathcal{O}_{\bullet}$, and vice versa.

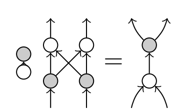
We will assume that the scalar  is always cancelable.

Proposition 3.3: In **FHilb** *if $O_{\circ}$ and $O_{\bullet}$ are self-adjoint operators correspoding to complementary observables, one can always choose a pair of coherent observable structures $(\mathcal{O}_{\circ}, \mathcal{O}_{\bullet})$ whose classical points correspond to the eigenbases of $O_{\circ}$ and $O_{\bullet}$.*

Proof: (sketch) The eigenbasis of a non-degenerate self-adjoint operator is only determined up to global phases. For a pair of mutually unbiased bases, it is always possible to choose these phases such that coherence is satisfied. ■

For this reason we will from now on assume that pairs of complementary observables are always coherent.

Definition 3.4: A pair $(\mathcal{O}_{\circ}, \mathcal{O}_{\bullet})$ of observables on the same object X is *strongly complementary* iff they are coherent and:

 (8)

Viewing one observable as monoid and the other as comonoid, the properties of coherence and strong complementarity state that a strongly complementary pair $(\mathcal{O}_{\circ}, \mathcal{O}_{\bullet})$ form a *scaled bialgebra*; that is, the defining equations of a bialgebra [37] hold upto a scalar multiple.

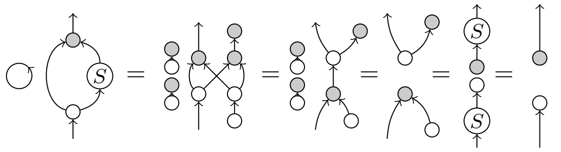
The following results about the antipode for a strongly complementary pair were shown in [28].

Lemma 3.5: Under the assumption that classical points are self-conjugate in their own colour, and provide 'enough points', the antipode S is self-adjoint, and is a Frobenius algebra endomorphism in both colours.

In fact, we can go further.

Theorem 3.6: Strong complementarity ⇒ complementarity.

Proof:



■

As a consequence, strongly complementary observables always form a *scaled Hopf algebra*. Note that Theorem 3.6 relies on the fact that both the monoid and the comonoid form a Frobenius algebra; it is certainly not the case that every scaled bialgebra is a Hopf algebra.

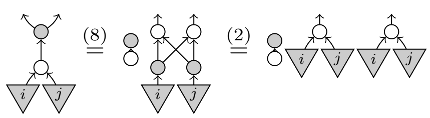
The converse to Theorem 3.6 does not hold: it is possible to find coherent complementary observables in **FHilb** which are not strongly complementary. See [9] for a counterexample.

A. Strong Complementarity and Phase Groups

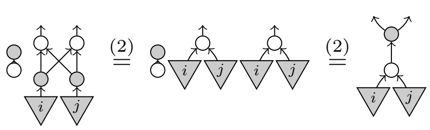
For complementary observables, classical points of one observable are always included in the phase group of the other observable, up to a normalizing scalar. Strong complementarity strengthens this property to inclusion as a subgroup. Let $\mathcal{K}_\circledcirc$ be the set of classical points of $\mathcal{O}_\circledcirc$ multiplied by the scalar factor .

Theorem 3.7: Let $(\mathcal{O}_\circ, \mathcal{O}_\circledcirc)$ be strongly complementary observables and let $\mathcal{O}_\circledcirc$ have finitely many classical points. Then $\mathcal{K}_\circledcirc$ forms a subgroup of the phase group $\Phi_\circ$ of $\mathcal{O}_\circ$. The converse also holds when $\mathcal{O}_\circledcirc$ has 'enough classical points'.

Proof: By strong complementarity it straightforwardly follows that, up to a scalar, $\mu_\circ$ applied to two classical points of $\mathcal{O}_\circledcirc$ yields again a classical point of $\mathcal{O}_\circledcirc$:



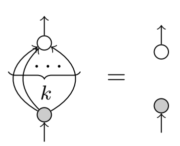
The unit of $\Phi_\circ$ is, up to a scalar, also a classical point of $\mathcal{O}_\circledcirc$ by coherence. Hence, $\mathcal{K}_\circledcirc$ is a submonoid of $\Phi_\circ$ and any finite submonoid is a subgroup. The converse follows from:



together with the 'enough classical points' assumption. ■

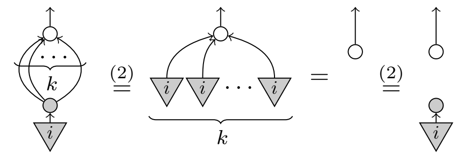
Recall that the exponent of a group G is the maximum order of any element of that group: $\exp(G) = \max\{|g| : g \in G\}$.

Corollary 3.8: For any pair of strongly complementary observables, let $k = \exp(\mathcal{K}_\circledcirc)$. Then, assuming $\mathcal{O}_\circledcirc$ has 'enough classical points':



(9)

Proof: In a finite abelian group, the order of any element divides $\exp(\mathcal{K}_\circledcirc)$. The result then follows by:



together with the 'enough classical points' assumption. ■

Proposition 3.9: For a pair of strongly complementary observables (i) is a $\mathcal{O}_\circledcirc$-homomorphism for all $(i) \in \mathcal{K}_\circledcirc$. Conversely, this property defines strong complementarity provided $\delta_\circ$ has 'enough classical points'.

Proof: Similar to the proof of Thm. 3.7. ■

B. Classification of Strong Complementarity in FHilb

Corollary 3.10: Every pair of strongly complementary observables in **FHilb** *is of the following form:*

$$\begin{cases} \delta_\circledcirc :: |g\rangle \mapsto |g\rangle \otimes |g\rangle \\ \epsilon_\circledcirc :: |g\rangle \mapsto 1 \end{cases} \qquad \begin{cases} \delta_\circ^\dagger :: |g\rangle \otimes |h\rangle \mapsto \frac{1}{\sqrt{D}}|g+h\rangle \\ \epsilon_\circ^\dagger :: 1 \mapsto \sqrt{D}|0\rangle \end{cases}$$

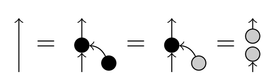
where $(G = \{g, h, \ldots\}, +, 0)$ is a finite Abelian group. Conversely, each such pair is always strongly complementary.

Proof: By Theorem 3.7 it follows that the classical points of one observable (here $\mathcal{O}_\circledcirc$) form a group for the multiplication of the other observable (here $\delta_\circ^\dagger$), and in **FHilb** this characterises strong complementarity. ■

One of the longest-standing open problems in quantum information is the characterisation of the number of pairwise complementary observables in a Hilbert space of dimension D. In all known cases this is $D+1$, and the smallest unknown case is $D = 6$. We now show that in the case of strong complementarity this number is always 2 for $D \geq 2$.

Theorem 3.11: In a Hilbert space with $D \geq 2$ the largest set of pairwise strongly complementary observables has size at most 2.

Proof: Assume that both $(\mathcal{O}_\circ, \mathcal{O}_\circledcirc)$ and $(\mathcal{O}_\circ, \mathcal{O}_\bullet)$ are strongly complementary pairs. By coherence and must be proportional to classical points of $\mathcal{O}_\circ$. If $(\mathcal{O}_\circledcirc, \mathcal{O}_\bullet)$ were to be strongly complementary observables, it is easily shown that $\neq 0$ so and are proportional to the same classical point. Hence, up to a non-zero scalar:

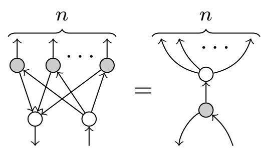


i.e. the identity has rank 1, which fails for $D \geq 2$. By Corollary 3.10 a strongly complementary pair exists for any $D \geq 2$. ■

IV. Diagrammatic Computation of GHZ Measurement Outcome Distributions

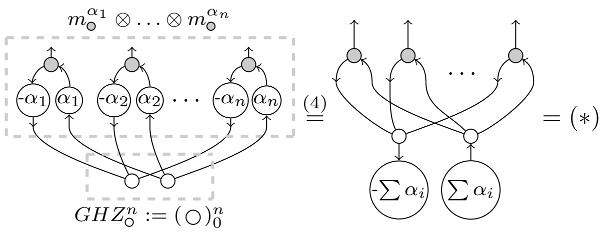
In order to present a graphical Mermin/GHZ style arguement, we show how to compute measurement outcomes for an n-party GHZ state graphically. This computation relies crucially on the following corollary, which follows from strong complementarity via a standard theorem about bialgebras.

Corollary 4.1: The following equation holds for any connected bipartite graph with directions as shown.

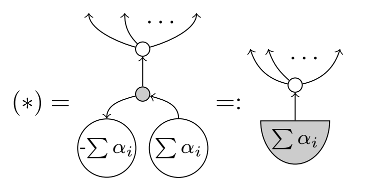


(10)

We compute the classical probability distributions (= $\mathcal{O}_\circ$-data) for n measurements against arbitrary phases $\alpha_i \in \Phi_\circ$ on n systems of any type in a generalised $GHZ^n_\circ$-state:



Applying Corollary 4.1, we note that this is a probability distribution followed by a $\circ$-copy.



(11)

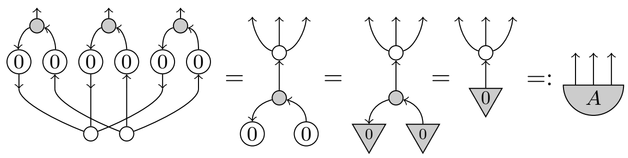
The following is an immediate consequence.

Theorem 4.2: When measuring each system of a GHZ^n_A-state against an arbitrary angle then the resulting classical probability distribution over outcomes is symmetric.

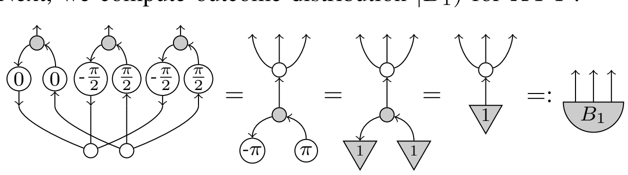
Theorem 4.3: The classical probability distributions for $m^{\alpha_1}_\circ \otimes \ldots \otimes m^{\alpha_n}_\circ$-measurements on a GHZ^n_A-state is:

- *uncorrelated if $|\sum \alpha_i)$ is a classical point for $\mathcal{O}_\circ$ and,*
- *parity-correlated if $|\sum \alpha_i)$ is a classical point i for $\mathcal{O}_\circ$ (i.e. contains precisely those outcomes $i_1 \otimes \ldots \otimes i_n$ such that the sum of group elements $\sum i_k$ is equal to i).*

Example 4.4: We can compute the outcome distributions for XXX, XYY, YXY, and YYX measurements on three qubits in a GHZ-state using the technique described above. First, outcome distribution $|A)$ for XXX:



Next, we compute outcome distribution $|B_1)$ for XYY:



Clearly, the other two cases, YXY and YYX, will give the same result. We therefore set $|B_1) = |B_2) = |B_3)$.

V. Mermin's Non-locality Argument in CQM

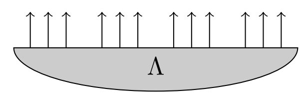
For a particular n-party state $|\Psi\rangle$ in some theory, a *local hidden variable (LHV)* model for that state consists of:

- a family of hidden states $|\lambda\rangle$, each of which assigns for any measurement on each subsystem a definite outcome,
- and, a probability distribution on these hidden states,

which simulates the probabilities of that theory. We say that a theory is *local* if each state admits a LHV model.

The Mermin argument is an important thought experiment which rules out the possibility that the predictions of quantum mechanics could be explained by LHV models [29]. Unlike Bell's argument, which merely shows that there exist quantum mechanical states whose outcome *probabilities* are inconsistent with locality assumptions, the Mermin argument demonstrates how there exists a quantum state whose outcome *possibilities* are inconsistent with locality.

Consider three systems and four possible (compound) measurement settings, consisting of the *control* XXX, and three *variations* XYY, YXY, and YYX. Let:



be a Born vector for $\mathcal{O}_\circ$ which represents the probability distribution on possible outcomes for each of these settings. This Born vector is a probability distribution over 'hidden states' each of which specifies the outcomes of all three systems for each of the four measurement settings, for example:

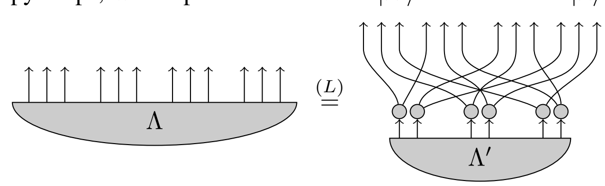
$$|\lambda\rangle = |\ \underbrace{+--}_{XXX}\ \underbrace{+++}_{XYY}\ \underbrace{--+}_{YXY}\ \underbrace{-+-}_{YYX}\ \rangle$$

yields outcomes $(+1,-1,-1)$ when XXX is measured, outcomes $(+1,+1,+1)$ when XYY is measured.

Assume now that this Born vector arises from an underlying LHV model (L). A hidden state of the LHV model stores one measurement outcome for each setting on each system:

$$|\lambda'\rangle = |\ \underbrace{\overbrace{+}^{X}\overbrace{-}^{Y}}_{system\ 1}\ \underbrace{\overbrace{-}^{X}\overbrace{+}^{Y}}_{system\ 2}\ \underbrace{\overbrace{-}^{X}\overbrace{+}^{Y}}_{system\ 3}\ \rangle$$

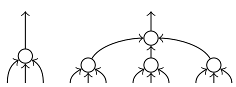
We can represent a probability distribution over such hidden states as a Born vector $|\Lambda')$, from which we generate $|\Lambda)$ via copy maps, which put a hidden state $|\lambda'\rangle$ into the form $|\lambda\rangle$:



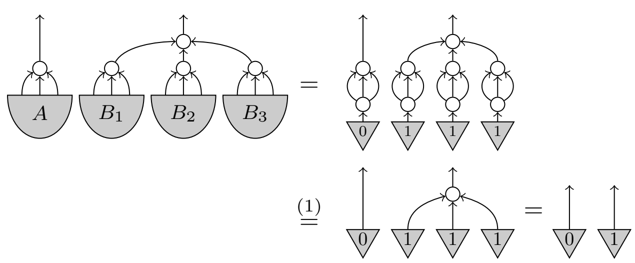
For example, the leftmost wire coming from $|\Lambda')$ represents the outcome when the first qubit of each hidden state is measured in X. This outcome is copied and sent to the first qubit for setting XXX and the first qubit for setting XYY.

We now investigate whether measurement of XXX, XYY, YXY and YYX on the qubits in a GHZ state is *consistent* (C) with such a LHV model. Although it is impossible to perform all four measurement setups simultaneously on the three qubits in the GHZ state, it should at least be true that any possible hidden state in $|\Lambda)$ must not be ruled out by the combined experimental data.

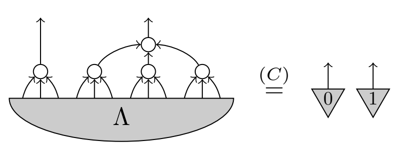
Since we are assuming that $|A)\otimes|B_1)\otimes|B_2)\otimes|B_3)$ is the probability distribution yielded by sampling $|\Lambda)$ independently for each measurement setup, it is a course-graining of $|\Lambda)$ itself. In other words, the support of $|\Lambda)$ must be contained in the support of $|A)\otimes|B_1)\otimes|B_2)\otimes|B_3)$.[2] In particular, if we can apply some function on hidden states in $|A)\otimes|B_1)\otimes|B_2)\otimes|B_3)$ that yields a definite outcome, that function must yield the same outcome in $|\Lambda)$. Consider the function:



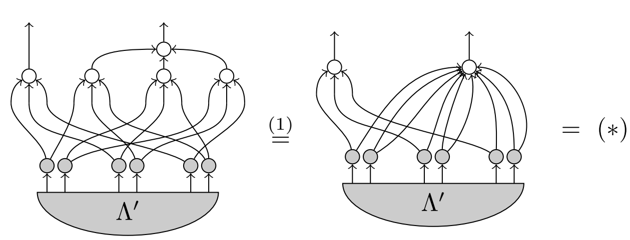
This function computes the parity (i.e. the $\mathbb{Z}_2$-sum) of outcomes for each of the four experiments. It then furthermore computes the *overall* parity of outcomes for the 3 variation experiments. We can straightforwardly show this yields a constant outcome on $|A)\otimes|B_1)\otimes|B_2)\otimes|B_3)$:



The last equation follows because by Theorem 3.7 the classical points for $\mathcal{O}_\circledcirc$ form a subgroup of the phase group $\Phi_\circ$. Since there are two classical points for $\mathcal{O}_\circledcirc$, this group must be $\mathbb{Z}_2$.

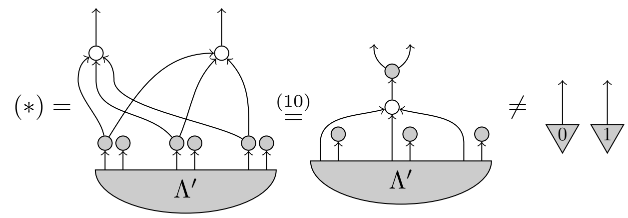


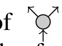
We can then use equations (C) and (L) to derive a contradiction. First, substitute (L) into the LHS of (C):



Since the group associated with $\circ$ is $\mathbb{Z}_2$, Cor. 3.8 applies for $k = 2$. Thus, we can delete pairs of parallel edges connecting dots of different colours. Then,

[2]Note that this is a statement about *possibilities* rather than *probabilities*.

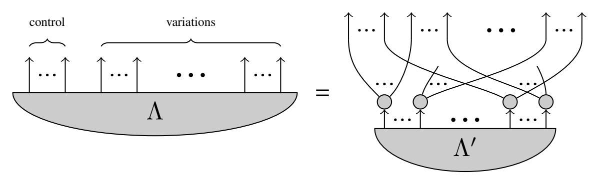


The final inequation follows from the distinctness of classical points and the cocommutativity of . Note that the symmetry from the LHS arises from the fact that after the application of Cor. 3.8, only the X measurements in the variations contribute to the output of the function.

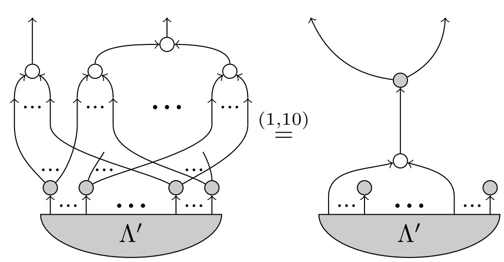
VI. Generalised Non-locality Arguments

To summarise, the conflict exposed in the previous section arises from the fact that experimental data, here from quantum theory, excludes a property imposed on the hidden states by locality. That property was symmetry. We now generalise this.

First we consider more general lists of measurement settings, represented again by a Born vector for $\mathcal{O}_\circledcirc$ generated from a LHV model:

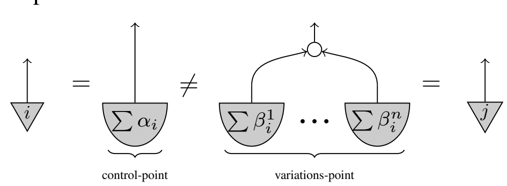


where we require, that via application of Cor. 3.8, we again obtain a symmetric expression:



This requires that for the variations, the multiplicity of occurrence of a particular measurement on the same system has to be a multiple of $exp(\mathcal{K}_\circledcirc)$.

Contradicting this symmetry requires an inequation between classical points:



Theorem 6.1: The above scenario provides a generalised Mermin non-locality argument whenever:

1) *the multiplicity of occurrence of measurements on the same system in the variations is a multiple of* $exp(\mathcal{K}_\circledcirc)$*,*
2) *and, the control-point and the variations-point are distinct classical points.*

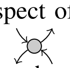
This theorem yields a wide variety of generalised Mermin-type non-locality arguments, and also characterises situations where such an argument fails to hold.

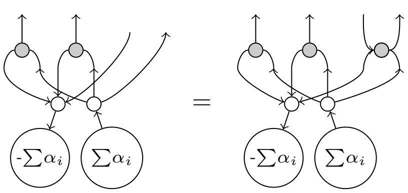
VII. Necessity of Strong Complementarity

In this section, we make an argument that the assumptions of the Mermin argument necessitate the use of strongly complementary observables. Consider a fixed 3-party GHZ-state $(\bigcirc)_0^3$. We will attempt to construct a Mermin argument by measuring all three systems with respect to some observables O_1, O_2, O_3. We make three assumptions about O_i.

1) **Phase-related:** All three O_i are within a $\mathcal{O}_\circ$-phase of some fixed observable structure $\mathcal{O}_\circledcirc$.
2) **Coherence:** $\mathcal{O}_\circledcirc$ is coherent with respect to $\mathcal{O}_\circ$.
3) **Sharpness:** After performing two of the three measurements on the GHZ-state, the remaining system is in an eigenstate of the third measurement.

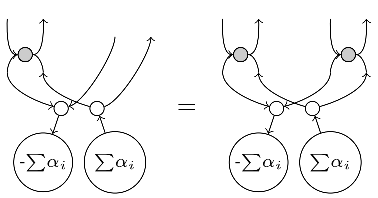
The first condition is satisfied by the usual Mermin argument, and can be seen as requiring that measurements differ in the 'maximally non-local' manner, since $\mathcal{O}_\circ$-phases can pass freely through the GHZ state. In **FHilb**, when $\mathcal{O}_\circledcirc$ is already complementary, we can choose it to be coherent by 3.3. The third assumption was highlighted by Mermin in [29] as an important aspect of the experimental setup.

The map  is called the *decoherence map* for $\mathcal{O}_\circledcirc$. It projects from the space of all quantum mixed states to the the space of classical mixtures of eigenstates of $\mathcal{O}_\circledcirc$. To assert sharpness, we require that, once two of the three systems are measured, the third is invariant under this map:



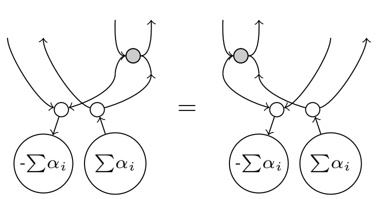
(12)

Plugging the unit of $\mathcal{O}_\circ$ in the 2nd system both for LHS and RHS, and using coherence we obtain:



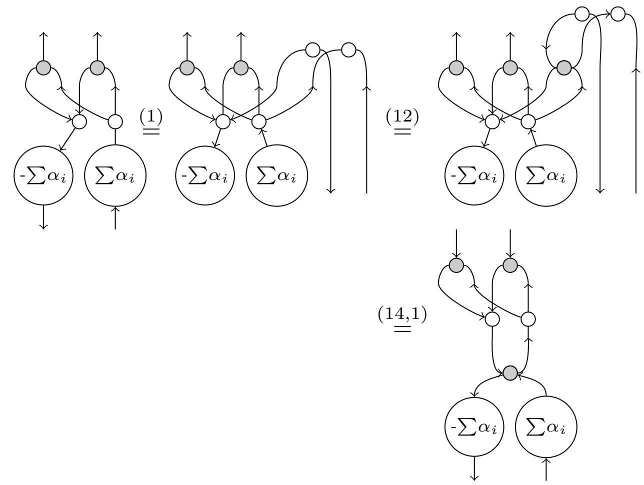
(13)

and by exploiting symmetry we have:



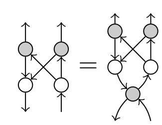
(14)

Hence we obtain:



Since $\delta_\circ^\dagger \circ (1_X \otimes \sum_i \alpha_i)$ is unitary it cancels.

Proposition 7.1: For a pair $(\mathcal{O}_\circ, \mathcal{O}_\circledcirc)$ of coherent observables on the same object the following equation implies (8):



(15)

Thus, for any coherent pair of observables, the sharpness of the third outcome necessitates strong complementarity.

VIII. Conclusion and Outlook

We cast Mermin's non-locality argument within CQM. This enabled us to substantially generalise this result to multiple parties, arbitrary angles and systems of arbitrary dimension.

The tools of CQM made most computations very easy as compared to the methods of standard quantum theory. We only rely on two simple rules: (i) contraction of labeled nodes (cf. (4)); (ii) commutation of the multiplications and comultiplications associated with distinctly colored nodes (cf. (10)).

In the graphical language, the manner in which the phases associated to the measurements interact through the GHZ-state effectively *depicts non-locality*.

The concepts required for reproducing Mermin-style non-locality arguments as well as the derivations crucially relied on the newly introduced notion of strong complementarity. We provided an operational interpretation for strong complementarity and classified strongly complementary observables in the case of quantum theory.

Our analysis also provides new insights in Mermin's argument. For example, while we crucially rely on 'strong' complementarity between the observable that characterises the GHZ state and the one for which the classical points form a subgroup of the corresponding phase group, complementarity between the X and Y obsevables is not at all essential when passing to the general case.

The results in this paper moreover open the door for a research program that concentrates on studying general process theories, in order to better understand what is so peculiar about the correlations encountered in quantum theory. CQM indeed provides an ideal arena for relating key concepts of quantum theory, and investigating in which manner they survive when passing to more general process theories.

The results in this paper may also have more direct applications to quantum computing. In particular, quantum secret sharing (QSS) [25] is a protocol which very closely resembles Mermin's non-locality argument. It relies on X or Y measurements on a qubit in a GHZ state, and the resulting correlations. Clearly the results in this paper would enable one to substantially generalise this schemes in a similar manner that we generalised Mermin's non-locality argument. Also, [24] contains initial investigations for a CQM-treatment of a wide range of quantum informatic protocols, which could be given a similar treatment and may lead to more generalisations of quantum communication schemes.

ACKNOWLEDGEMENT

We acknowledge support from US Office of Naval Research (ONR), Foundational Questions Institute (FQXi), *Fonds de la Recherche Scientifique* (FNRS), and John Templeton Foundation (JTF).

REFERENCES

[1] S. Abramsky (2009) *No-Cloning In Categorical Quantum Mechanics*arxiv:arXiv:0910.2401v1.

[2] S. Abramsky and B. Coecke (2004) *A categorical semantics of quantum protocols*. In: Proceedings of 19th IEEE conference on Logic in Computer Science, pages 415–425. IEEE Press. arXiv:quant-ph/0402130. Revised version (2009): *Categorical quantum mechanics*. In: Handbook of Quantum Logic and Quantum Structures, K. Engesser, D. M. Gabbay and D. Lehmann (eds), pages 261–323. Elsevier. arXiv:0808.1023

[3] H. Barnum, J. Barrett, L. O. Clark, M. Leifer, R. W. Spekkens, N. Stepanik, A. Wilce and R. Wilke (2010) *Entropy and information causality in general probabilistic theories*. New Journal of Physics **12**, 033024. arXiv:0909.5075

[4] J. Barrett (2007) *Information processing in generalized probabilistic theories.* Physical Review A **75**, 032304.

[5] C. H. Bennett and G. Brassard (1984) *Quantum cryptography: Public key distribution and coin tossing*. In *Proceedings of IEEE International Conference on Computers, Systems, and Signal Processing*, pages 175–179, Bangalore, India, 1984. IEEE, New York.

[6] H. R. Brown and G. Svetlichny (1990) *Nonlocality and Gleason's lemma. Part I. Deterministic theories*Foundations of Physics **20**, 1379–1387.

[7] G. Chiribella, G. M. D'Ariano and P. Perinotti (2010) *Informational derivation of quantum theory*. Physical Review A **84**, 012311. arXiv:1011.6451

[8] B. Coecke and R. Duncan (2008) *Interacting quantum observables*. In: Proceedings of the 35th International Colloquium on Automata, Languages and Programming (ICALP), pp. 298–310, Lecture Notes in Computer Science 5126, Springer-Verlag. Extended version (2011):

[9] B. Coecke and R. Duncan (2011) *Interacting quantum observables: categorical algebra and diagrammatics*. New Journal of Physics **13**, 043016. arXiv:0906.4725

[10] B. Coecke and B. Edwards (2011) *Toy quantum categories*. In: Proceedings of Quantum Physics and Logic/Development of Computational Models (QPL-DCM). Electronic Notes in Theoretical Computer Science **271** (1), 29–40. arXiv:0808.1037

[11] B. Coecke, B. Edwards and R. W. Spekkens (2011) *Phase groups and the origin of non-locality for qubits*. Electronic Notes in Theoretical Computer Science **271** (2), 15–36. arXiv:1003.5005

[12] B. Coecke and A. Kissinger (2010) *The compositional structure of multipartite quantum entanglement*. In: Proceedings of the 37th International Colloquium on Automata, Languages and Programming (ICALP), pp. 297–308, Lecture Notes in Computer Science 6199, Springer-Verlag. arXiv:1002.2540

[13] B. Coecke, E. O. Paquette and D. Pavlovic (2009) *Classical and quantum structuralism*. In: Semantic Techniques for Quantum Computation, I. Mackie and S. Gay (eds), pages 29–69, Cambridge University Press. arXiv:0904.1997

[14] B. Coecke and D. Pavlovic (2007) *Quantum measurements without sums*. In: Mathematics of Quantum Computing and Technology, G. Chen, L. Kauffman and S. Lamonaco (eds), pages 567–604. Taylor and Francis. arXiv:quant-ph/0608035

[15] B. Coecke, D. Pavlovic and J. Vicary (2008) *A new description of orthogonal bases*. To appear in Mathematical Structures in Computer Science. arXiv:0810.0812

[16] D. G. B. J. Dieks (1982) *Communication by EPR devices*. Physics Letters A **92**, 271-272.

[17] R. Duncan and S. Perdrix (2010) *Rewriting measurement-based quantum computations with generalised flow*. In: Proceedings of the 37th International Colloquium on Automata, Languages and Programming (ICALP), Lecture Notes in Computer Science 6199, Springer-Verlag.

[18] B. Edwards (2009) *Non-locality in Categorical Quantum Mechanics*. DPhil thesis, University of Oxford.

[19] B. Edwards (2010) *Phase groups and local hidden variables*. Research Report RR-10-15, Dept. of Computer Science, University of Oxford.

[20] A. Ekert (1991) *Quantum cryptography based on Bell's theorem*. Physical Review Letters **67**, 661–663.

[21] D. M. Greenberger, M. A. Horne, A. Shimony and A. Zeilinger (1990) *Bell's theorem without inequalities*. American Journal of Physics **58**, 1131–1143.

[22] L. Hardy (2011) *Foliable operational structures for general probabilistic theories*. In: Deep Beauty: Understanding the Quantum World through Mathematical Innovation. H. Halvorson (ed.), pages 409–442, Cambridge University Press. arXiv:0912.4740

[23] P. Heywood and M. L. G. Redhead (1983) *Nonlocality and the Kochen-Specker paradox*. Foundations of Physics **13**, 481–499.

[24] A. Hillebrand (2011) *Quantum protocols involving multiparticle entanglement and their representations in the zx-calculus*. MSc. thesis, University of Oxford. http://www.cs.ox.ac.uk/people/bob.coecke/Anne.pdf

[25] M. Hillery, V. Bužek, and A. Berthiaume (1999) *Quantum secret sharing*. Physical Review A **59**, 1829–1834.

[26] C. Horsman (2011) *Quantum picturalism for topological cluster-state computing*. New Journal of Physics **13** 095011. arXiv:1101.4722

[27] G. M. Kelly and M. L. Laplaza (1980) *Coherence for compact closed categories*. Journal of Pure and Applied Algebra **19**, 193–213.

[28] A. Kissinger (2012) *Pictures of Processes*. DPhil thesis, University of Oxford.

[29] N. D. Mermin (1990) *Quantum mysteries revisited*. American Journal of Physics **58**, 731–734.

[30] A. K. Pati and S. L. Braunstein (2000) *Impossibility of deleting an unknown quantum state*. Nature **404**, 164–165. arXiv:quant-ph/9911090

[31] D. Pavlovic (2009) *Quantum and classical structures in nondeterminstic computation*. Lecture Notes in Computer Science **5494**, page 143–157, Springer. arXiv:0812.2266

[32] M. Pawlowski, T. Paterek, D. Kazlikowski, V. Scarani, A. Winter and M. Zukowski (2009) *Information causality as a physical principle*. Nature **461**, 1101. arXiv:0905.2292

[33] E. Schrödinger (1935) *Discussion of probability relations between separated systems*. Proceedings of the Cambridge Philosophical Society **31**, 555–563.

[34] P. Selinger (2007) *Dagger compact closed categories and completely positive maps*. Electronic Notes in Theoretical Computer Science **170**, pages 139–163.

[35] P. Selinger (2011) *A survey of graphical languages for monoidal categories*. In: New Structures for Physics, B. Coecke (ed.), pages 289–356. Lecture Notes in Physics 813, Springer-Verlag. arXiv:0908.3347

[36] R. W. Spekkens (2007) *Evidence for the epistemic view of quantum states: A toy theory*. Physical Review A **75**, 032110. arXiv:quant-ph/0401052

[37] R. Street (2007) *Quantum Groups: A Path to Current Algebra*. Cambridge University Press.

[38] W. K. Wootters and W. Zurek (1982) *A single quantum cannot be cloned*. Nature **299**, 802–803.